

# METHOD AND SYSTEM FOR OPTIMALLY SELECTING A WEB FIREWALL IN A TCP/IP NETWORK

## *Technical field of the invention*

5 The present invention relates to computer networks, and more particularly to a method and system in a TCP/IP network for optimally selecting a Web Firewall according to some response time and availability criteria.

## *Background art*

### **INTERNET**

10 The Internet is a global network of computers and computers networks (the "Net"). The Internet connects computers that use a variety of different operating systems or languages, including UNIX, DOS, Windows, Macintosh, and others. To facilitate and allow the communication among these various systems and languages, the Internet uses a <sup>protocol</sup>~~language~~ referred to as TCP/IP ("Transmission Control Protocol/Internet Protocol"). TCP/IP protocol supports three basic applications on the Internet :

- transmitting and receiving electronic mail,
- 20 • logging into remote computers (the "Telnet"), and
- transferring files and programs from one computer to another ("FTP" or "File Transfer Protocol").

### **WORLD WIDE WEB**

25 With the increasing size and complexity of the Internet, tools have been developed to help find information on the network, often called navigators or navigation systems. Navigation systems that have been developed include standards

such as Archie, Gopher and WAIS. The World Wide Web ("WWW" or "the Web") is a recent superior navigation system. The Web is:

- an Internet-based navigation system,
- an information distribution and management system for the Internet, and
- a dynamic format for communicating on the Web.

The Web seamlessly, for the <sup>user</sup> ~~use~~, integrates <sup>the formatting</sup> ~~format~~ of information, including still images, text, audio and video. A user on the Web using a graphical user interface ("GUI", pronounced "gooey") may transparently communicate with different host computers on the system, and different system applications (including FTP and Telnet), and different information formats for files and documents including, for example, text, <sup>audio</sup> ~~sound~~ and <sup>image (graphics)</sup> ~~graphics~~.

## HYPERMEDIA

The Web uses hypertext and hypermedia. Hypertext is a subset of hypermedia and refers to computer-based "documents" in which readers move from one place to another in a document, or to another document, in a non-linear manner. To do this, the Web uses a client-server architecture. The Web servers enable the user to access hypertext and hypermedia information through the Web and the user's computer. (The user's computer is referred to as a client computer of the Web Server computers.) The clients send requests to the Web Servers, which react, search and respond. The Web allows client application software to request and receive hypermedia documents (including formatted text, audio, video and graphics) with hypertext link capabilities to other hypermedia documents, from a Web file server.

The Web, then, can be viewed as a collection of document files residing on Web host computers that are interconnected by hyperlinks using networking protocols, forming a virtual "web" that spans the Internet.

## UNIFORM RESOURCE LOCATORS

A resource of the Internet is unambiguously identified by a Uniform Resource Locator (URL), which is a pointer to a particular resource at a particular location. A URL specifies the protocol used to access a server (e.g. HTTP, FTP,...), the name of the server, and the location of a file on that server.

## HYPER TEXT TRANSFER PROTOCOL

Each Web page that appears on client monitors of the Web may appear as a complex document that integrates, for example, text, images, sounds and animation. Each such page may also contain hyperlinks to other Web documents so that a user at a client computer using a mouse may click on icons and may activate hyperlink jumps to a new page (which is a graphical representation of another document file) on the same or a different Web server.

A Web server is a software program on a Web host computer that answers requests from Web clients, typically over the Internet. All Web<sup>Servers</sup> use a language or protocol to communicate with Web clients which is called Hyper Text Transfer Protocol ("HTTP"). All types of data can be exchanged among Web servers and clients using this protocol, including Hyper Text Markup Language ("HTML"), graphics, sound and video. HTML describes the layout, contents and hyperlinks of the documents and pages. Web clients when browsing :

- convert user specified commands into HTTP GET requests,
- connect to the appropriate Web server to get information, and
- wait for a response. The response from the server can be the requested document or an error message.

After the document or an error message is returned, the connection between the Web client and the Web server is closed.

*The first*

*a* ~~First~~, version of HTTP is a stateless protocol. That is with HTTP, there is no continuous connection between each client and each server. The Web client using HTTP receives a response as HTML data or other data. This description applies to version 1.0 of HTTP protocol, while the ~~new version~~ *newer versions* 1.1 *breaks* ~~break~~ this barrier of stateless protocol by keeping the connection between the server and client alive under certain conditions.

## **BROWSER**

*a* After receipt, the Web client formats and presents the data or activates an ancillary application such <sup>as</sup> a sound player to present the data. To do this, the server or the client determines the various types of data received. The Web Client is also referred to as the Web Browser, since it in fact browses documents retrieved from the Web Server.

## **DOMAIN NAMES**

*a* The host or computers names (like www.entreprise.com) are translated into numeric Internet addresses (like 194.56.78.3), and vice versa, by using a method called DNS ("Domain Name Service"). DNS is supported by network-resident servers, also known as domain name servers or DNS servers.

## **INTRANET**

*a* Some companies use the same mechanism as the Web to communicate inside their own corporation. In this case, this mechanism is called an "Intranet". These companies use the same networking/transport protocols and locally based Web servers to provide access to vast amount of corporate information in a cohesive fashion. As this data may be private to the corporation, and because the members of the company still need to have access to public Web information, to avoid that people not belonging to the company can access to this private Intranet coming from the public Internet, they protect

the access to their network by using a special equipment called a Firewall.

## **FIREWALL**

5 A Firewall protects one or more computers with Internet connections from access by external computers connected to the Internet. A Firewall is a network configuration, usually created by hardware and software, that forms a boundary between networked computers within the Firewall from those outside the Firewall. The computers within the Firewall form a  
10 secure sub-network with internal access capabilities and shared resources not available from the outside computers.

Often, a single machine, on which the Firewall is, allows access to both internal and external computers. Since the computer, on which the Firewall is, directly interacts with  
15 the Internet, strict security measures against unwanted access from external computers are required.

A Firewall is commonly used to protect information such as electronic mail and data files within a physical building or organization site. A Firewall reduces the risk of intrusion by unauthorized people from the Internet, however, the same  
20 security measures can limit or require special software for those inside the Firewall who wish to access information on the outside. A Firewall can be configured using "Proxies" or "Socks" to designate access to information from each side of  
25 the Firewall.

## **PROXY SERVER**

A HTTP Proxy is a special server that typically runs in conjunction with Firewall software and allows an access to the Internet from within a Firewall. The Proxy Server :

- 30
- waits for a request (for example a HTTP request) from inside the Firewall,
  - forwards the request to the remote server outside the Firewall,

- reads the response, and
- sends the response back to the client.

5 A single computer can run multiple servers, each server connection identified with a port number. A Proxy Server, like an HTTP Server or a FTP Server, occupies a port. Typically, a connection uses standardized port numbers for each protocol (for example, HTTP = 80 and FTP = 21). That is why an end user has to select a specific port number for each defined Proxy Server. Web Browsers usually let the end user set the host name and port number of the Proxy Servers in a customizable panel. Protocols such as HTTP, FTP, Gopher, WAIS, and Security can usually have designated Proxies. Proxies are generally preferred over Socks for their ability to perform caching, high-level logging, and access control, because they provide a specific connection for each network service protocol.

## SOCKS

20 <sup>a</sup>Socks Server (also called <sup>a</sup>Socks Gateway) is also ~~a~~ software that allows computers inside a Firewall to gain access to the Internet. Socks is usually installed on a server positioned either inside or on the Firewall. Computers within the Firewall access the Socks Server as clients to reach the Internet. Web Browsers usually let the end user set the host name and port number of the Socks hosts (servers) in a customizable panel. On some Operating Systems, the host is specified in a separate file (e.g. socks.conf file). As the Socks Server acts <sup>as</sup> ~~a~~ a layer underneath the protocols (HTTP, FTP, ..), it cannot cache data (as <sup>a</sup>Proxy does), because it doesn't decode the protocol to know what kind of data ~~/~~ it is transferring.

## 30 OPTIONS

<sup>that</sup>  
The Web Browser often proposes <sup>a</sup>the end user ~~to~~ select between the different options <sup>of</sup> "No Proxies", "Manual Proxy

Configuration", or "Automatic Proxy Configuration" to designate the conduit between his computer and the Internet.

- Users with a direct connection to the Internet should use the default, which is "No Proxies".
- If the Intranet is protected by one or several Firewalls, the end user may :
  - select one of these Firewalls as the elected Proxy, by entering its host name into the "Manual Proxy Configuration", or
  - automatically refer~~s~~<sup>Proxy</sup> to the enterprise policy in terms of ~~proxies~~<sup>A</sup> attribution between locations, by pointing to a common configuration file in a remote server. This is done by choosing the "Automatic Proxy Configuration" and by providing the Web Browser with the unique address of the common configuration file ("Universal Resource Locator" or "URL") located in the remote server.

Today, most of the Web Browsers are configured to forward all requests, even requests for internal hosts, through the Socks Firewall. So when the end user wants to have access to an internal Web-based application, his request travels to the Firewall, and is then reflected back into the internal network. This sends internal traffic on a long path, puts extra load on the Firewall and on the network, and worst of all, slows down the response time the end user sees from the applications and Web pages he is trying to access. This is called "non flexible" Socks access (when everything goes via the Socks Server).

#### MANUAL PROXY CONFIGURATION

The Manual Proxy configuration in the Web Browser is simple to process, but its main drawback is that the Firewall (or Proxy) selection is then static. There is no dynamic criterion for the Firewall selection, such as selection of the

Firewall providing the best response time. Firewall failures require a manual reconfiguration of the navigation software to point to another active Firewall, since the manual configuration usually only allows the definition of one single Firewall per protocol with no possibility to pre-configure a backup Firewall. In addition to the manual proxy configuration in the Web Browser, external procedures can be used to provide some kind of robustness in the Firewall selection. They rely for instance on the use of multiple Firewalls having the same name defined as aliases in the Domain Name Server (DNS). But this technique based on alias definition still has drawbacks since for instance the DNS is not always contacted for name resolution by Web Clients caching locally the name resolution. Other techniques using external hardware equipment such as load and request dispatcher provide more robustness and load balancing, but still have drawbacks such as the need for additional and costly hardware.

#### **AUTOMATIC PROXY CONFIGURATION**

Automatic Proxy Configuration (or also referred to as "autoproxy") can set the location of the HTTP, FTP, and Gopher Proxy every time the Web Browser is started. An autoproxy retrieves a file of address ranges and instructs the Web Browser to either directly access internal IBM hosts or to go to the Socks Server to access hosts on the Internet.

Automatic Proxy Configuration is more desirable than simple Proxy Server Configuration in the Web Browser, because much more sophisticated rules can be implemented about the way Web pages are retrieved (directly or indirectly). Automatic Proxy Configuration is useful to users, because the Web Browser knows how to retrieve pages directly if the Proxy Server fails. Also Proxy requests can be directed to another or multiple Proxy Servers at the discretion of the system administrator, without the end user <sup>having</sup> ~~has~~ to make any additional changes to his Web Browser configuration. In general, these Proxy configuration files (also called autoproxy code) are



a usually written in <sup>the</sup> Javascript language. Autoproxy facility can also contain a file of address ranges for instructing the Web Browser to either directly access internal hosts or to go to the Socks Server to access hosts on the Internet. The Socks Server protects the internal network from unwanted public access while permitting access of network members to the Internet. One of the drawbacks of this "autoproxy" mechanism is that there is no proactive Firewall failure detection nor response time consideration.

5 More explanations about the domain presented in the above sections can be found in the following publications ~~incorporated herewith by reference:~~

- "Java Network Programming" by Elliotte Rusty Harold, published by O'Reilly, February 1997.
- "Internet in a nutshell" by Valerie Quercia, published by O'Reilly, October 1997.
- "Building Internet Firewalls" by Brent Chapman and Elizabeth Zwickky, published by O'Reilly, September 1995.

## PROBLEM

20 The problem to solve is to provide an optimized Web access, with a dynamic Proxy or Socks Server selection to get the best response time, and a detection of failures in Proxy or Socks Server to prevent Web service disruption. The current solutions address this problem partially:

- 25 • Web Browsers can be manually configured with the target Proxy or Socks Server. The main drawbacks of this solution are the following :
  - There is no dynamic Proxy/Socks Server selection. A manual reconfiguration of the Web Browser upon Proxy/Socks Server failure is required.
  - 30 • Only a "manual" load balancing through the Web Browser static configuration is provided.

- Proxy/Socks Server names must be known and manually configured by end users.
- Web Browsers can be configured with their autoproxy feature, using a static list of target Proxy/Socks Servers downloaded from a dedicated autoproxy URL (Uniform Resource Locator) system. The main drawback of this solution is the following:
  - There is no response time consideration in the Proxy/Socks Server selection, nor efficient Proxy/Socks Server failure detection (i.e. Web Browser waits for time-out before switching to backup, even at initial autoproxy loading) in the Proxy/Socks Server selection.

An alternate to these current solutions is to cluster the Proxy/Socks Servers using an external dispatcher system acting as single logical access point. All Web Browsers are then manually configured with the name of that external dispatcher system (as the target Proxy/Socks Server) which then routes the traffic to a selected Proxy/Socks Server. An example of such a dispatcher is for example the IBM Interactive Network Dispatcher product. More information concerning this product can be found in IBM's publication entitled "Interactive Network Dispatcher V1.2 - User's Guide" GC31-8496-01 ~~incorporated herewith by reference.~~ Although a dispatcher oriented solution allows an efficient load balancing in most cases, its main drawback is that additional dedicated system or specific hardware is required, and that the external dispatcher name has to be manually configured by the end users in their Web Browsers.

### ***Objects of the invention***

- The object of the present invention is to optimize Proxy/Socks Server selection by using availability and response time criteria.
- 5 • It is a further object of the present invention to optimize the Web service performance by integrating a response time factor to the Proxy/Socks Server selection.
- It is another object of the present invention to minimize Web service interruption and thus to insure a better service availability by automatically detecting Proxy/Socks Server~~s~~ failures.

### ***Summary of the invention***

The present invention relates to dynamic autoproxy configuration and more particularly to a method and system for optimizing selection of a Proxy/Socks Server according to some response time and availability criteria. The invention rests on a dynamic autoproxy mechanism using availability and response time probes.

The present invention also relies on probes retrieving well known HTML pages through each Proxy/Socks Server, measuring associated response time, detecting Proxy/Socks failures and degradation of response time.

The present invention also uses a CGI (Common Gateway Interface) program for dynamically creating autoproxy code (in a preferred embodiment Javascript code) on an autoproxy URL (Universal Resource Locator) system for selecting the Proxy/Socks Server using availability and response time information provided by probes.

The present invention fixes the drawbacks of the existing current solutions by integrating dynamic Proxy/Socks availability and response time selection criteria to the autoproxy mechanism.

5 The present inventions provides the following advantages:

- Early detection of Proxy/Socks Servers failures provides a high Web service availability.
- Integration of a response time factor to the Proxy/Socks Server selection optimizes the Web service performances.
- 10 • Induced HTTP survey traffic is minimized by running availability and response time probes from a single autoproxy URL system (compared with running the probes on each Web Browser system).
- Integration of response time degradation in the probes achieves a proactive Proxy/Socks Servers failure detection.
- Periodical dynamic update of "best" Proxy/Socks Server can be provided to Web Browser.
- Useless traffic to failing Proxy/Socks Server is minimized since Proxy/Socks Servers are excluded from list of available target servers upon failure detection.
- 20 • No additional or specific hardware is required.
- Ease of Web Browser configuration provided to mobile users (Web Browser is configured once).
- a • Web Browser performances <sup>is</sup> ~~are~~ not degraded because  
25 availability and response time probes are not processed within the downloaded autoproxy code (Javascript code) but in the autoproxy URL system.

### ***Drawings***

30 The novel and inventive features believed characteristics of the invention are set forth in the appended claims. The

invention itself, however, as well as a preferred mode of use, further objects and advantages thereof, will best be understood by reference to the following detailed description of an illustrative detailed embodiment when read in conjunction with the accompanying drawings, wherein :

- Figure 1 is a general logical view of an end user system interfacing a Web Browser for accessing the World Wide Web according to prior art.
- Figure 2 is a general physical view of the set-up shown in Figure 1, according to prior art.
- Figure 3 is a logical view of the availability and response time probes external flows according to the present invention.
- Figure 4 is a flow chart showing the internal logic flow of the availability and response time probe introduced in Figure 3 according the present invention.
- Figure 5 is a physical view of the logical environment described in Figure 3 according to the present invention.
- Figure 6 is a view of the data flows associated with the entities depicted in Figure 5, according to the present invention.
- Figure 7 depicts the storage of the availability and response times probes measurements, according to the present invention.
- Figure 8 is a flow chart of the program running on the autoproxxy URL (Universal Resource Locator) system, according to the present invention.

### ***Preferred embodiment of the invention***

The present invention relies on dynamic autoproxy configuration and more particularly to a method and system for selecting a Proxy/Socks Server according to some response time and availability criteria. It rests on a dynamic autoproxy mechanism using availability and response time probes. It relies on probes retrieving well known HTML pages through each Proxy/Socks Server, measuring associated response time, detecting Proxy/Socks failures and degradation of response time.

It also uses a CGI (Common Gateway Interface) program for dynamically creating autoproxy code (in a preferred embodiment Javascript code) on an autoproxy URL (Universal resource locator) system for selecting said Proxy/Socks Server.

### **LOGICAL VIEW OF A END USER ACCESSING THE WORLD WIDE WEB**

Figure 1 shows a user system with a user interface (102) running a program known as a Web Browser (101) which enables access to the World-Wide-Web (WWW). The WWW content is transferred using the HTTP protocol. HTTP requests and responses are going to and from the Web Browser program (101) and a destination Web Server (103) containing the WWW content the user tries to access. The Firewall (104) between the Web Browser (101) and the Web Server (103) acts as an intermediary HTTP Proxy forwarding the HTTP requests and responses to their destination. The Web Browser program (101) makes an HTTP request to the Firewall (104) and the Firewall forwards the request to the destination Web Server (103). The flow in the reverse direction is the HTTP response which again goes via the Firewall (104) on its way to the Web Browser (101). In this way the Firewall can limit the traffic to the transactions it is configured to allow (based on some defined

a security and access control policy). The Firewall hence protects the network where <sup>the</sup> Web Browser is located.

#### GENERAL PHYSICAL VIEW OF AN END USER ACCESSING THE WWW

5 Figure 2 is a physical view of the set-up shown logically in Figure 1. In this particular example, the Web Browser (201) runs on a system attached to an Intranet (202). The Firewalls (203) that protect the Intranet attach both the (private) Intranet (202) and the (public) Internet (204). <sup>The</sup> ~~the~~ destination Web Server (205) also connects to the Internet. 10 This is the environment where the Web Browser, Firewalls, and Web Server perform their function when the user is "browsing" the Internet WWW. It is important to note the fact the Firewalls attach two networks and hence are able to act as the intermediary for communications between the two networks. Multiple Firewalls are often uses in order to provide some degree of access robustness and load sharing.

#### LOGICAL VIEW OF AVAILABILITY AND RESPONSE TIME PROBES

20 The domain of the invention is the one described in Figure 1 and Figure 2, where a user within an Intranet wants to access the World-Wide-Web using a Web Browser, and where the Intranet network is protected from the Internet by several Firewalls playing role of so called HTTP Proxies (Figure 2). The issue is to select the "best" Proxy/Socks Server to insure an optimized availability and response time of the service to the end user. To automatically optimize this selection, a software component called a "WWW availability and response time probe" is introduced. Its role is to provide selection criteria. As shown in Figure 3, this data is gathered by measuring the response time for requesting a specific content of a well known Web Server. The induced HTTP survey traffic is 30 minimized, by running the availability and response time probes from single autoproxy URL system (versus running the probes on each Web Browser client system).

Figure 3 demonstrates the function of a flexible WWW availability and response time probe and the way it can be used to gather measurements on the availability and response time of both HTTP Proxies and Socks Servers. The upper part of Figure 3 details the interaction of the probe with an HTTP Proxy Server (304). The client system (302) that runs the probe (configured to test proxies) basically requests Web content (page) from the Web Server (307) via the Proxy Server (304) similar to the process shown in Figure 1. The HTTP request in this case represents an "HTTP survey flow" (303) to the Proxy Server. The Proxy Server forwards (306) the request to the Web Server (via the Firewall (305) which is not depicted). The client system times how long the request/response HTTP survey flow takes and uses this information as a measurement of the response time and availability via the tested Proxy Server (for the Web content samples that were tested). If the client system is also the autoproxy URL system (301) then this measurement information for each Proxy Server can be used to work out a sense of the "best" Proxy Server to use. This can then be encoded in the autoproxy URL that the Web Browser programs use to work out their correct Proxy Server to use.

The lower portion of the Figure 3 shows a similar arrangement but in this case the measurements data is being gathered for a Socks Server (Gateway) access method. Again a client probe (309) makes an HTTP request that represents an "HTTP survey flow" (310) which travels via the Socks Server (311) and then onto (312) the destination Web site (313). This HTTP request is for a set target URL (308) that is known to exist on the target Web Server. Again it is the timing of how long this survey takes that provides the measurements data that can be used to generate an autoproxy URL that takes into account the relative performance of a set of Socks Server (or in the case above, HTTP Proxies).

Obviously if there is no response to the HTTP survey flow, then the particular Proxy or Socks Server being tested



can be marked as unavailable. In this way the autoproxy URL can be used to not select Proxy or Socks Servers that do not work .

#### INTERNAL LOGIC OF THE AVAILABILITY AND RESPONSE TIME PROBES

5       The internal mechanism of the probe itself is described in Figure 4. The probe simulates a Web Client, by requesting through an HTTP connection a Web page from a target URL through the target Proxy/Socks Server (using its host name and port as a reference). The Web page is retrieved either through a normal HTTP connection, or through a socksified flow (a flow through a Socks Server). Typically, <sup>a</sup> a normal flow is used to retrieve a Web page from a Proxy Server or from a Web Server, while socksified flow is used to retrieve a Web page through a Socks Server. Then, the probe basically checks that the Web page :

- is received within an allowed amount of time in seconds, and
- contains a specific keyword to make sure that the received page is correct.

When these two conditions are fulfilled, the Web page retrieval is successful.

25       Finally, the probe returns either the associated response time in seconds (successful retrieval) or a failure return code. This mechanism retrieves one or multiple target Web pages. When multiple Web pages are retrieved, the probe program sequentially tests each Web page until one Web page retrieval is successful or all Web page retrievals fail. Probes :

- retrieve well known HTML pages through each Proxy/Socks Server,
- measure associated response time, and also

- detect Proxy/Socks Server failures and response time degradation.

Figure 4 is a flow chart showing the internal logic flow of the WWW availability and response time probe introduced in Figure 3.

- The first thing the probe program does is to start a timer (401).
- Next the probe program attempts to establish a connection (402) with the target Web Server to retrieve a Web page at the target URL (Universal Resource locator). The probe program establishes the connection according to the way it has been configured e.g. via an HTTP Proxy Server, via a Socks Server (Gateway) or directly.
- If the attempt for establishing the connection is unsuccessful, the probe program immediately goes into error mode (408). An error value is returned (407) by the probe program indicating that the connection is not possible.
- If the attempt for establishing the connection is successful, then the Web page (403) is retrieved by the probe program.
- The probe program then closes the connection (404) pursuant to the normal HTTP protocol procedure.
- To ensure that the Web page has been correctly retrieved, the probe program then searches for known keywords (405) that are expected to be in the Web page.
- If the keyword is found (406) in the Web page, then the Web page retrieval is successful. The timer is stopped and the correct response time for the operation is returned. By storing and integrating a short historic of the measured response time over time, the probe program can detect and return any response time degradation, thus enabling an anticipation of the Proxy/Socks Servers failures.

- a 5
- If however the correct keyword is not found (407) in the Web page, then the Web retrieval is unsuccessful and again an error value is returned. The type of event that might trigger this sort of error is when the connection is successfully established but <sup>a</sup> Web page with an error is retrieved.
  - The action whereby the probe goes into retry mode (409) occurs only when the probe is configured to try multiple destination URL's as opposed to a single URL. This adds some robustness to the testing of the probe and hence insulates it somewhat from one-off network "glitches" (e.g. dropped connections etc.).
- 10

#### PHYSICAL VIEW OF AVAILABILITY AND RESPONSE TIME PROBES EXTERNAL FLOWS

005 20 006 20 007 20 008 20 009 20 010 20 011 20 012 20 013 20 014 20 015 20 016 20 017 20 018 20 019 20 020 20 021 20 022 20 023 20 024 20 025 20 026 20 027 20 028 20 029 20 030 20 031 20 032 20 033 20 034 20 035 20 036 20 037 20 038 20 039 20 040 20 041 20 042 20 043 20 044 20 045 20 046 20 047 20 048 20 049 20 050 20 051 20 052 20 053 20 054 20 055 20 056 20 057 20 058 20 059 20 060 20 061 20 062 20 063 20 064 20 065 20 066 20 067 20 068 20 069 20 070 20 071 20 072 20 073 20 074 20 075 20 076 20 077 20 078 20 079 20 080 20 081 20 082 20 083 20 084 20 085 20 086 20 087 20 088 20 089 20 090 20 091 20 092 20 093 20 094 20 095 20 096 20 097 20 098 20 099 20 100 20 101 20 102 20 103 20 104 20 105 20 106 20 107 20 108 20 109 20 110 20 111 20 112 20 113 20 114 20 115 20 116 20 117 20 118 20 119 20 120 20 121 20 122 20 123 20 124 20 125 20 126 20 127 20 128 20 129 20 130 20 131 20 132 20 133 20 134 20 135 20 136 20 137 20 138 20 139 20 140 20 141 20 142 20 143 20 144 20 145 20 146 20 147 20 148 20 149 20 150 20 151 20 152 20 153 20 154 20 155 20 156 20 157 20 158 20 159 20 160 20 161 20 162 20 163 20 164 20 165 20 166 20 167 20 168 20 169 20 170 20 171 20 172 20 173 20 174 20 175 20 176 20 177 20 178 20 179 20 180 20 181 20 182 20 183 20 184 20 185 20 186 20 187 20 188 20 189 20 190 20 191 20 192 20 193 20 194 20 195 20 196 20 197 20 198 20 199 20 200 20 201 20 202 20 203 20 204 20 205 20 206 20 207 20 208 20 209 20 210 20 211 20 212 20 213 20 214 20 215 20 216 20 217 20 218 20 219 20 220 20 221 20 222 20 223 20 224 20 225 20 226 20 227 20 228 20 229 20 230 20 231 20 232 20 233 20 234 20 235 20 236 20 237 20 238 20 239 20 240 20 241 20 242 20 243 20 244 20 245 20 246 20 247 20 248 20 249 20 250 20 251 20 252 20 253 20 254 20 255 20 256 20 257 20 258 20 259 20 260 20 261 20 262 20 263 20 264 20 265 20 266 20 267 20 268 20 269 20 270 20 271 20 272 20 273 20 274 20 275 20 276 20 277 20 278 20 279 20 280 20 281 20 282 20 283 20 284 20 285 20 286 20 287 20 288 20 289 20 290 20 291 20 292 20 293 20 294 20 295 20 296 20 297 20 298 20 299 20 300 20 301 20 302 20 303 20 304 20 305 20 306 20 307 20 308 20 309 20 310 20 311 20 312 20 313 20 314 20 315 20 316 20 317 20 318 20 319 20 320 20 321 20 322 20 323 20 324 20 325 20 326 20 327 20 328 20 329 20 330 20 331 20 332 20 333 20 334 20 335 20 336 20 337 20 338 20 339 20 340 20 341 20 342 20 343 20 344 20 345 20 346 20 347 20 348 20 349 20 350 20 351 20 352 20 353 20 354 20 355 20 356 20 357 20 358 20 359 20 360 20 361 20 362 20 363 20 364 20 365 20 366 20 367 20 368 20 369 20 370 20 371 20 372 20 373 20 374 20 375 20 376 20 377 20 378 20 379 20 380 20 381 20 382 20 383 20 384 20 385 20 386 20 387 20 388 20 389 20 390 20 391 20 392 20 393 20 394 20 395 20 396 20 397 20 398 20 399 20 400 20 401 20 402 20 403 20 404 20 405 20 406 20 407 20 408 20 409 20 410 20 411 20 412 20 413 20 414 20 415 20 416 20 417 20 418 20 419 20 420 20 421 20 422 20 423 20 424 20 425 20 426 20 427 20 428 20 429 20 430 20 431 20 432 20 433 20 434 20 435 20 436 20 437 20 438 20 439 20 440 20 441 20 442 20 443 20 444 20 445 20 446 20 447 20 448 20 449 20 450 20 451 20 452 20 453 20 454 20 455 20 456 20 457 20 458 20 459 20 460 20 461 20 462 20 463 20 464 20 465 20 466 20 467 20 468 20 469 20 470 20 471 20 472 20 473 20 474 20 475 20 476 20 477 20 478 20 479 20 480 20 481 20 482 20 483 20 484 20 485 20 486 20 487 20 488 20 489 20 490 20 491 20 492 20 493 20 494 20 495 20 496 20 497 20 498 20 499 20 500 20 501 20 502 20 503 20 504 20 505 20 506 20 507 20 508 20 509 20 510 20 511 20 512 20 513 20 514 20 515 20 516 20 517 20 518 20 519 20 520 20 521 20 522 20 523 20 524 20 525 20 526 20 527 20 528 20 529 20 530 20 531 20 532 20 533 20 534 20 535 20 536 20 537 20 538 20 539 20 540 20 541 20 542 20 543 20 544 20 545 20 546 20 547 20 548 20 549 20 550 20 551 20 552 20 553 20 554 20 555 20 556 20 557 20 558 20 559 20 560 20 561 20 562 20 563 20 564 20 565 20 566 20 567 20 568 20 569 20 570 20 571 20 572 20 573 20 574 20 575 20 576 20 577 20 578 20 579 20 580 20 581 20 582 20 583 20 584 20 585 20 586 20 587 20 588 20 589 20 590 20 591 20 592 20 593 20 594 20 595 20 596 20 597 20 598 20 599 20 600 20 601 20 602 20 603 20 604 20 605 20 606 20 607 20 608 20 609 20 610 20 611 20 612 20 613 20 614 20 615 20 616 20 617 20 618 20 619 20 620 20 621 20 622 20 623 20 624 20 625 20 626 20 627 20 628 20 629 20 630 20 631 20 632 20 633 20 634 20 635 20 636 20 637 20 638 20 639 20 640 20 641 20 642 20 643 20 644 20 645 20 646 20 647 20 648 20 649 20 650 20 651 20 652 20 653 20 654 20 655 20 656 20 657 20 658 20 659 20 660 20 661 20 662 20 663 20 664 20 665 20 666 20 667 20 668 20 669 20 670 20 671 20 672 20 673 20 674 20 675 20 676 20 677 20 678 20 679 20 680 20 681 20 682 20 683 20 684 20 685 20 686 20 687 20 688 20 689 20 690 20 691 20 692 20 693 20 694 20 695 20 696 20 697 20 698 20 699 20 700 20 701 20 702 20 703 20 704 20 705 20 706 20 707 20 708 20 709 20 710 20 711 20 712 20 713 20 714 20 715 20 716 20 717 20 718 20 719 20 720 20 721 20 722 20 723 20 724 20 725 20 726 20 727 20 728 20 729 20 730 20 731 20 732 20 733 20 734 20 735 20 736 20 737 20 738 20 739 20 740 20 741 20 742 20 743 20 744 20 745 20 746 20 747 20 748 20 749 20 750 20 751 20 752 20 753 20 754 20 755 20 756 20 757 20 758 20 759 20 760 20 761 20 762 20 763 20 764 20 765 20 766 20 767 20 768 20 769 20 770 20 771 20 772 20 773 20 774 20 775 20 776 20 777 20 778 20 779 20 780 20 781 20 782 20 783 20 784 20 785 20 786 20 787 20 788 20 789 20 790 20 791 20 792 20 793 20 794 20 795 20 796 20 797 20 798 20 799 20 800 20 801 20 802 20 803 20 804 20 805 20 806 20 807 20 808 20 809 20 810 20 811 20 812 20 813 20 814 20 815 20 816 20 817 20 818 20 819 20 820 20 821 20 822 20 823 20 824 20 825 20 826 20 827 20 828 20 829 20 830 20 831 20 832 20 833 20 834 20 835 20 836 20 837 20 838 20 839 20 840 20 841 20 842 20 843 20 844 20 845 20 846 20 847 20 848 20 849 20 850 20 851 20 852 20 853 20 854 20 855 20 856 20 857 20 858 20 859 20 860 20 861 20 862 20 863 20 864 20 865 20 866 20 867 20 868 20 869 20 870 20 871 20 872 20 873 20 874 20 875 20 876 20 877 20 878 20 879 20 880 20 881 20 882 20 883 20 884 20 885 20 886 20 887 20 888 20 889 20 890 20 891 20 892 20 893 20 894 20 895 20 896 20 897 20 898 20 899 20 900 20 901 20 902 20 903 20 904 20 905 20 906 20 907 20 908 20 909 20 910 20 911 20 912 20 913 20 914 20 915 20 916 20 917 20 918 20 919 20 920 20 921 20 922 20 923 20 924 20 925 20 926 20 927 20 928 20 929 20 930 20 931 20 932 20 933 20 934 20 935 20 936 20 937 20 938 20 939 20 940 20 941 20 942 20 943 20 944 20 945 20 946 20 947 20 948 20 949 20 950 20 951 20 952 20 953 20 954 20 955 20 956 20 957 20 958 20 959 20 960 20 961 20 962 20 963 20 964 20 965 20 966 20 967 20 968 20 969 20 970 20 971 20 972 20 973 20 974 20 975 20 976 20 977 20 978 20 979 20 980 20 981 20 982 20 983 20 984 20 985 20 986 20 987 20 988 20 989 20 990 20 991 20 992 20 993 20 994 20 995 20 996 20 997 20 998 20 999 20 1000 20 1001 20 1002 20 1003 20 1004 20 1005 20 1006 20 1007 20 1008 20 1009 20 1010 20 1011 20 1012 20 1013 20 1014 20 1015 20 1016 20 1017 20 1018 20 1019 20 1020 20 1021 20 1022 20 1023 20 1024 20 1025 20 1026 20 1027 20 1028 20 1029 20 1030 20 1031 20 1032 20 1033 20 1034 20 1035 20 1036 20 1037 20 1038 20 1039 20 1040 20 1041 20 1042 20 1043 20 1044 20 1045 20 1046 20 1047 20 1048 20 1049 20 1050 20 1051 20 1052 20 1053 20 1054 20 1055 20 1056 20 1057 20 1058 20 1059 20 1060 20 1061 20 1062 20 1063 20 1064 20 1065 20 1066 20 1067 20 1068 20 1069 20 1070 20 1071 20 1072 20 1073 20 1074 20 1075 20 1076 20 1077 20 1078 20 1079 20 1080 20 1081 20 1082 20 1083 20 1084 20 1085 20 1086 20 1087 20 1088 20 1089 20 1090 20 1091 20 1092 20 1093 20 1094 20 1095 20 1096 20 1097 20 1098 20 1099 20 1100 20 1101 20 1102 20 1103 20 1104 20 1105 20 1106 20 1107 20 1108 20 1109 20 1110 20 1111 20 1112 20 1113 20 1114 20 1115 20 1116 20 1117 20 1118 20 1119 20 1120 20 1121 20 1122 20 1123 20 1124 20 1125 20 1126 20 1127 20 1128 20 1129 20 1130 20 1131 20 1132 20 1133 20 1134 20 1135 20 1136 20 1137 20 1138 20 1139 20 1140 20 1141 20 1142 20 1143 20 1144 20 1145 20 1146 20 1147 20 1148 20 1149 20 1150 20 1151 20 1152 20 1153 20 1154 20 1155 20 1156 20 1157 20 1158 20 1159 20 1160 20 1161 20 1162 20 1163 20 1164 20 1165 20 1166 20 1167 20 1168 20 1169 20 1170 20 1171 20 1172 20 1173 20 1174 20 1175 20 1176 20 1177 20 1178 20 1179 20 1180 20 1181 20 1182 20 1183 20 1184 20 1185 20 1186 20 1187 20 1188 20 1189 20 1190 20 1191 20 1192 20 1193 20 1194 20 1195 20 1196 20 1197 20 1198 20 1199 20 1200 20 1201 20 1202 20 1203 20 1204 20 1205 20 1206 20 1207 20 1208 20 1209 20 1210 20 1211 20 1212 20 1213 20 1214 20 1215 20 1216 20 1217 20 1218 20 1219 20 1220 20 1221 20 1222 20 1223 20 1224 20 1225 20 1226 20 1227 20 1228 20 1229 20 1230 20 1231 20 1232 20 1233 20 1234 20 1235 20 1236 20 1237 20 1238 20 1239 20 1240 20 1241 20 1242 20 1243 20 1244 20 1245 20 1246 20 1247 20 1248 20 1249 20 1250 20 1251 20 1252 20 1253 20 1254 20 1255 20 1256 20 1257 20 1258 20 1259 20 1260 20 1261 20 1262 20 1263 20 1264 20 1265 20 1266 20 1267 20 1268 20 1269 20 1270 20 1271 20 1272 20 1273 20 1274 20 1275 20 1276 20 1277 20 1278 20 1279 20 1280 20 1281 20 1282 20 1283 20 1284 20 1285 20 1286 20 1287 20 1288 20 1289 20 1290 20 1291 20 1292 20 1293 20 1294 20 1295 20 1296 20 1297 20 1298 20 1299 20 1300 20 1301 20 1302 20 1303 20 1304 20 1305 20 1306 20 1307 20 1308 20 1309 20 1310 20 1311 20 1312 20 1313 20 1314 20 1315 20 1316 20 1317 20 1318 20 1319 20 1320 20 1321 20 1322 20 1323 20 1324 20 1325 20 1326 20 1327 20 1328 20 1329 20 1330 20 1331 20 1332 20 1333 20 1334 20 1335 20 1336 20 1337 20 1338 20 1339 20 1340 20 1341 20 1342 20 1343 20 1344 20 1345 20 1346 20 1347 20 1348 20 1349 20 1350 20 1351 20 1352 20 1353 20 1354 20 1355 20 1356 20 1357 20 1358 20 1359 20 1360 20 1361 20 1362 20 1363 20 1364 20 1365 20 1366 20 1367 20 1368 20 1369 20 1370 20 1371 20 1372 20 1373 20 1374 20 1375 20 1376 20 1377 20 1378 20 1379 20 1380 20 1381 20 1382 20 1383 20 1384 20 1385 20 1386 20 1387 20 1388 20 1389 20 1390 20 1391 20 1392 20 1393 20 1394 20 1395 20 1396 20 1397 20 1398 20 1399 20 1400 20 1401 20 1402 20 1403 20 1404 20 1405 20 1406 20 1407 20 1408 20 1409 20 1410 20 1411 20 1412 20 1413 20 1414 20 1415 20 1416 20 1417 20 1418 20 1419 20 1420 20 1421 20 1422 20 1423 20 1424 20 1425 20 1426 20 1427 20 1428 20 1429 20 1430 20 1431 20 1432 20 1433 20 1434 20 1435 20 1436 20 1437 20 1438 20 1439 20 1440 20 1441 20 1442 20 1443 20 1444 20 1445 20 1446 20 1447 20 1448 20 1449 20 1450 20 1451 20 1452 20 1453 20 1454 20 1455 20 1456 20 1457 20 1458 20 1459 20 1460 20 1461 20 1462 20 1463 20 1464 20 1465 20 1466 20 1467 20 1468 20 1469 20 1470 20 1471 20 1472 20 1473 20 1474 20 1475 20 1476 20 1477 20 1478 20 1479 20 1480 20 1481 20 1482 20 1483 20 1484 20 1485 20 1486 20 1487 20 1488 20 1489 20 1490 20 1491 20 1492 20 1493 20 1494 20 1495 20 1496 20 1497 20 1498 20 1499 20 1500 20 1501 20 1502 20 1503 20 1504 20 1505 20 1506 20 1507 20 1508 20 1509 20 1510 20 1511 20 1512 20 1513 20 1514 20 1515 20 1516 20 1517 20 1518 20 1519 20 1520 20 1521 20 1522 20 1523 20 1524 20 1525 20 1526 20 1527 20 1528 20 1529 20 1530 20 1531 20 1532 20 1533 20 1534 20 1535 20 1536 20 1537 20 1538 20 1539 20 1540 20 1541 20 1542 20 1543 20 1544 20 1545 20 1546 20 1547 20 1548 20 1549 20 1550 20 1551 20 1552 20 1553 20 1554 20 1555 20 1556 20 1557 20 1558 20 1559 20 1560 20 1561 20 1562 20 1563 20 1564 20 1565 20 1566 20 1567 20 1568 20 1569 20 1570 20 1571 20 1572 20 1573 20 1574 20 1575 20 1576 20 1577 20 1578 20 1579 20 1580 20 1581 20 1582 20 1583 20 1584 20 1585 20 1586 20 1587 20 1588 20 1589 20 1590 20 1591 20 1592 20 1593 20 1594 20 1595 20 1596 20 1597 20 1598 20 1599 20 1600 20 1601 20 1602 20 1603 20 1604 20 1605 20 1606 20 1607 20 1608 20 1609 20 1610 20 1611 20 1612 20 1613 20 1614 20 1615 20 1616 20 1617 20 1618 20 1619 20 1620 20 1621 20 1622 20 1623 20 1624 20 1625 20 1626 20 1627 20 1628 20 1629 20 1630 20 1631 20 1632 20 1633 20 1634 20 1635 20 1636 20 1637 20 1638 20 1639 20 1640 20 1641 20 1642 20 1

The use of response time and availability criteria also provides a proactive Proxy/Socks Servers failure detection through the integration of response time degradation. The Web Browser can be periodically and dynamically updated with a new selection of the "best" Proxy/Socks Server using :

- "refresh" tag in the autoprox code,
- external code (or Java applet), or
- a new feature in the Web Browser for periodically and automatically refreshing the autoprox code.

Another positive consequence is the minimization of the useless traffic to failing Proxy/Socks Server since Proxy/Socks Servers are excluded from list of available target servers upon failure detection. Since an autoprox mechanism is used, there is no need for manually updating the manual proxy configuration in the Web Browser in case of Proxy/Socks Server failure. Proxy/Socks Servers names or locations don't need to be known and configured by the end user, thus providing for instance a seamless service for mobile users.

Figure 5 is a physical view of the logical environment described in Figure 3. The Web Browser (501) attached to the Intranet (502) is configured to use an autoprox URL to determine which Proxy/Socks Server (Firewall) (503) to use for having access to the Internet (504) and the destination Web Server (505). The system where the autoprox URL resides (506) also runs the availability and response time probes (507) configured to test the Proxy/Socks Servers. The autoprox URL uses the CGI (Common Gateway Interface) (508) to dynamically generate the autoprox code of the autoprox URL. The autoprox code is based on the information gathered by the availability and response time probes. In this way the Web Browser is configured with

- an available Proxy/Socks Server, and

- what is deemed the best Proxy/Socks Server.

#### DATA FLOWS OF AVAILABILITY AND RESPONSE TIME PROBES

Figure 6 is a view of the actual data flows associated with the entities depicted in Figure 5. Again, the Web Browser (601) attached to the Intranet (602) is configured to use an autoproxxy URL to determine which Proxy/Socks Server (Firewall) (603) to use for having access to the Internet (604) and the destination Web Server (605). The Web Browser has access (609) to the autoproxxy URL system (606) to first determine which Proxy/Socks Server it should use. The Web Browser can be periodically and dynamically updated with the "best" Proxy/Socks Server using :

- the "refresh" tag in the autoproxxy code,
- an external code (or Java applet), or
- a Web Browser new feature for periodically and automatically refreshing the autoproxxy code.

The system where the autoproxxy URL resides and which runs the availability and response time probes (607), uses the CGI (Common Gateway Interface) (608) to dynamically generate the autoproxxy code (610) of the autoproxxy URL. The autoproxxy code is based on the information gathered by the availability and response time probes that have tested the Proxy/Socks Servers via the HTTP survey flows (611 and 612) described in Figure 3. In this way the Web Browser ends up with what is deemed the best Proxy/Socks Server.

#### INTERNAL STORAGE OF AVAILABILITY AND RESPONSE TIME PROBES

Figure 7 depicts the internal storage within the autoproxxy URL system of the information retrieved by the availability and response times probes (701). Each probe updates (702) a table in the autoproxxy URL system (703) with the measurements of each Proxy/Socks Server it tests. In this way the table contains the current state of all Proxy/Socks

Servers that are candidate to be selected and used by the Web Browser. At configurable or periodical time intervals, probes test again the Proxy/Socks Servers (704) and the cycle again is repeated.

## 5 PROGRAM RUNNING AT AUTOPROXY URL SYSTEM

Figure 8 again refers to the internal logic of the program running on the autoprox proxy URL system.

- The autoprox proxy URL system is initially contacted by a Web Browser (801) wanting to "know" which is the best Proxy/Socks Server (or Firewall) to use. This is for instance achieved by selecting the Automatic Proxy Configuration option in the Web Browser and by providing information such as the URL of the autoprox proxy code.
- The autoprox proxy URL system activates (802) the CGI (Common Gateway Interface) program (via the Web Server CGI extensions). The CGI program has access to all standard CGI variables including the IP (Internet Protocol) address of the Web Browser.
- The CGI program selects (803) the best Proxy/Socks Server for the client system (Web Browser) based on both the IP address of the Web Browser (obtained as a CGI variable) and the information generated by the availability and response time probes for each Proxy/Socks Server and stored in the table of the autoprox proxy URL system (807). The IP address is used to add a geographical criteria to the Proxy/Socks Server selection. For instance, if two Proxy/Socks Servers provide the same response time (one in the US, the other one in Europe), the closest Proxy/Socks Server is preferred (the one in Europe if the Web Browser is in Europe).
- To improve the robustness of the Proxy/Socks Server selection, the CGI program selects a best "backup" Proxy/Socks Server for the Web Browser (804). This "backup" Proxy/Socks Server is automatically used by the Web Browser after it times out attempting to use what it thinks is the

"best" Proxy/Socks Server. Again this "backup" Proxy/Socks Server is selected using both the IP address of the Web Browser (obtained as a CGI variable) and the information generated by the availability and response time probes for each Proxy/Socks Server and stored in the table of the autoproxxy URL system (807).

- Once the CGI program has selected the best and backup Proxy/Socks Servers, it created the autoproxxy code (805). This code is generally made of Javascript language.
- Once the autoproxxy code has been created, the autoproxxy URL system downloads it to the Web Browser (806) via standard HTTP protocol as any other output of a CGI program.

While the invention has been particularly shown and described with reference to a preferred embodiment, it will be understood that various changes in form and detail may be made therein without departing from the spirit, and scope of the invention.